AIR POLLUTION EXPOSURE AND LUNG FUNCTION IN CHILDREN: A MICRO-EPIDEMIOLOGICAL STUDY

by

Alan J. Krupnick
and
Winston Harrington

assisted by Tracy Narel

August 9, 1985

Resources for the Future
1616 P Street, N.W.
Washington, D.C. 20036

Submitted to the Economics Analysis Branch, Office of Policy Planning and Evaluation, USEPA. Contract No, 68-01-6596, Work Assignment S-041.
Project Officer: Bart Ostro.

DISCLAIMER

TABLE OF CONTENTS

Tables

Executive Summary

Few epidemiological studies have attempted to derive a functional relationship between lung function of children and daily air pollution concentrations. The present study derives concentration-response functions for the effects of daily average TSP, RSP, sulfate, and nitrate concentrations on the pulmonary function scores (FEV.75) of up to 4,800 elementary school children living in one of six Birmingham, Alabama communities. The children were tested for pulmonary function in each of three rounds (October-November 1972, January-February 1973, April-May 1973). Before testing began, a chronic disease questionnaire based on the British Medical Research Council's Chronic Bronchitis Questionnaire was sent home with each child to be filled out by the mother. It contained demographic, socioeconomic, and health-related questions, as well as a question on the presence of a gas stove.

Twenty-four hour average pollution concentrations at six monitors were collected by EPA for total suspended particulates (TSP), respirable particulates (RSP), sulfates and nitrates. Because RSP and TSP are so highly correlated, we dropped one of these measures (RSP) from further analysis.

Linear and non-linear regression was used to analyze, in each round, the relationship between FEV scores and age, height, sex, race (white or other), packs per day smoked by the mother, and, separately, by the father, an interaction term for age by sex by mother's smoking status, type of stove heat (gas or other), education of family head, number of family members per room, family size, mother head of household (yes or no), the presence of active asthma or cold symptoms on the test day, the presence of chronic bronchitis, and dummy variables for the teams and the spirometers used in the lung function tests. The latter variables were used to capture any biases in the administration of the tests or the measurement of test scores. Finally, we added pollution variables, including the 24-hour average concentrations of pollutants the day of the FEV exam and the day before the exam, these terms in squared or log form (to test for particular nonlinear relationships), the average daily values for the previous week, the maximum daily concentration for the previous week, and the average for the previous month.

We also estimated regressions for the subpopulation of normal, asymptomatic children and then the subpopulation of symptomatic children who participated in the Birmingham study. A symptomatic child is defined as one who had asthma symptoms or cold symptoms on the day of the test, or who has chronic bronchitis. An asymptomatic child is defined as one who has none of these.

Linear regressions with the full sample reveal that the variables in this general model explain, in each of the three rounds, about 74% of the variation in FEV scores. Age, height, sex, and race account for most of this variation, along with dummy variables for presence of chronic bronchitis, asthma symptoms, and cold symptoms. Of lesser importance, but with generally significant, positive effects on FEV is educational attainment. Family size and crowding in the house have unexpected positive, but small, effects on FEV scores. Both the team performing the tests and the spirometer used to measure the results were found to significantly affect FEV scores.

The TSP and sulfate pollution variables are found to take the expected negative sign and to have significant effects on FEV. The TSP coefficient is quite stable and significant in fall and spring; the sulfate coefficient is highly significant in all rounds although its size varies by a factor of 5 over the three seasons. The TSP elasticities (the percentage change in the mean FEV caused by a one percent change in the mean concentration of a pollutant) range from -0.004 to -0.019, with most in the -0.014 to -0.019 range. Even at the high end of this range the effects are quite small. Sulfates appear to have somewhat larger elasticities--from -0.077 to -0.072. The effect of a 50% increase in either TSP or sulfates would be equivalent to being about one year younger or (since age and height are correlated) one-seventh of an inch shorter.

By comparing the size of the comparable and significant pollution coefficients for the asymptomatic and symptomatic regressions, it is apparent that pollution has a greater effect on the FEV of symptomatic children than on the FEV of asymptomatic children, although the differences between comparable coefficients are significant at the 95% level only for sulfates in the fall round. Crowding, education, and family size are all less robust variables in the split samples. In addition, crowding, at least in the fall and winter rounds, appears to be more beneficial to symptomatics than asymptomatics.

Attempts to isolate the effect of lagged pollution exposures on FEV l scores were generally unsuccessful. In addition, alternative nonlinear, specifications of the lung function-air pollution relationship performed no better than the linear specification. We also found that daily pollution data are generally better at explaining variation in lung function scores than monthly average data.

Acknowledgements

We would like to thank Bart Ostro, our EPA project officer, for his insightful comments and suggestions, as well as his unflagging financial support. EPA's Victor Hasselblad was, as always, extremely helpful and prompt in providing us with the data analyzed in this study. During his stay as a Natural Science Fellow at Resources for the Future, Robert Frank, M.D. of the Department of Environmental Sciences at the Johns Hopkins University School of Hygiene and Public Health, gave us an understanding of the functioning of the lung and its relationship to respiratory disease. This background and his advice during the planning and conduct of this study improved the quality of our research immeasurably. Finally, we would like to thank Cheryl Johnson for her skill and perserverance in typing this paper.

Introduction

Few epidemiological studies have attempted to derive a functional relationship between lung function of children and daily TSP and SO_2 exposure (USEPA, 1982, pp. 14-44). Lunn (1967 and 1970) found significant associations between lung function (FEV.75) in children and exposures to SO_2 and particulate matter in ambient air, although it appears that communities were characterized by their yearly, rather than daily, pollution levels. Lawther, et al. (1970) used daily "British Smoke" and SO_2 measures to search for associations between air pollution and lung function. However, adults with bronchitis were the target group. Later work by Lawther, et al. (1973 and 1974) broadened the target group to include healthy adults, but data for only four healthy adults and two bronchitics were analyzed.

The present study derives concentration-response functions for the effects of daily average TSP, RSP, sulfate, and nitrate concentrations on lung function in school-age children living in one of six Birmingham, Alabama communities. Although this data set has been available since 1977, it has been analyzed only once--along with data gathered from twenty-four other communities in six additional SMSA's. Hasselblad, et al. (1981) used ANOVA to explain variation in FEV.75 with variables for passive smoke exposure, presence of gas stove, education of family head, and the child’s age, height, and sex. In addition, and of most relevance here, they found that community of residence as a proxy for pollution exposure and other environmental risks is a significant risk factor. While one may suspect that pollution differences across communities may explain some of the variation in FEV scores, attempts by Hasselblad, et al. to explicitly include a wide range of pollutant measures were generally unsatisfactory.

The present study explicitly examined the effect of daily and monthly pollution concentrations on lung function. Tests were also made of the effect on lung function of the variables found in Hasselblad, et al. and some additional variables--family size, race, crowding (number of rooms per family member), and having a mother as head of household. The effect of equipment and team member bias on lung function was also examined. The above tests were conducted on both the full sample and sub-samples of individuals with and without asthma, cold symptoms, or bronchitis. In addition, an analysis of lagged pollution effects was performed.[1] Finally, we compared regression results with daily pollution data to those with monthly average pollution data to test alternative hypotheses about the temporal relationship of air pollution to lung function.

Data Characteristics

The data for this study were collected in Birmingham, Alabama, from October 1972 to May 1973 as part of the EPA's Community Health Environmental Surveillance System (CHESS). Birmingham was chosen for the high level of particulates in some communities. Elementary school children (aged six to fourteen) within each of six communities were tested for pulmonary function (FEV.75) in each of three rounds (October-November 1972, January-February 1973, April-May 1973). Measurements were made with a 12-L bellows-type spirometers with digital readout[2]. Each child was given spirometer tests during one day in each round until three acceptable measurements were obtained. Then, the maximum reading for that round was

---

1. Dockery, et al. (1981) found that declines in pulmonary function persisted for one to three weeks after exposure to episodic levels of TSP and SO_2.

2. See Hasselblad, et al. (1981) for details on protocols and instrumentation.

chosen. Symptoms at the time of testing were also noted.

Before testing began, a chronic disease questionnaire based on the British Medical Research Council's Chronic Bronchitis Questionnaire was sent home with each child to be filled out by the mother. It contained demographic, socioeconomic, and health-related questions, as well as a question on the presence of a gas stove. Table 1 presents summary statistics for the non-pollution data used in this study. Only children living and attending school. within one mile of the same monitor participated in the study.

Twenty-four hour average pollution concentrations at six monitors were collected by EPA for total suspended particulates (TSP), respirable particulates (RSP), sulfates and nitrates. Table 2 provides the average and maximum ambient concentrations of pollutants by monitor by season for the observations in our sample, and a correlation matrix. Because RSP and TSP are so highly correlated, we dropped one of these measures (RSP) from further analysis. Because RSP data are missing for many days during the fall and spring rounds, using TSP instead of RSP also increases the sample sizes in these two rounds.

When the health data set was matched to the daily pollution data, it became apparent that the pollution data had a number of gaps, particularly in Fall 1972. Many observations would have been lost if we required each child to have complete data in all three periods. Accordingly, this restriction was relaxed and the number of observations in the regressions now varies somewhat by period, by pollutant mix, and by lag structure.

Analytical Methods

Linear and non-linear regression was used to analyze, in each round, the relationship between FEV scores and age, height, sex, race (white or

Table 1
Descriptive Statistics

	Fall (N = 1792)				Winter (N = 4303)				Spring (N = 4816)			
	Mean (%)	SD	Min	Max	Mean (%)	SD	Min	Max	Mean (%)	SD	Min	Max
FEV (liters)	1.42	0.39	.47	3.14	1.51	.40	.53	3.34	1.55	.40	.47	3.51
AGE1	9.03	1.80	6	13	10.00	1.73	7	14	10.02	1.74	7	14
HEIGHT (inches)	54.03	4.7	33	69	54.76	4.59	35.5	69.5	55.18	4.51	41	71
BOYS (%)	(52)				(51)				(51)			
WHITE (%)	(50)				(77)				(72)			
PPDMOM (avg)[1]	0.32				0.31				0.30			
PPDDAD (avg)	0.45				0.50				0.49			
GAS (%)[2]	(22)				(20)				(20)			
EDUC (years)[3]	11.70	3.06	6	18	12.45	2.89	6	18	12.24	2.95	6	18
CROWD[4]	1.16	0.47	0.17	4.5	1.30	0.45	0.17	4.5	1.26	0.45	0.17	4.5
FAMSIZE (#)	5.69	2.19	2	15	5.17	1.74	2	5	5.29	1.84	2	15
MOMHEAD[5] (%)	(20.0)				(12.3)				(13.4)			
BRONC[6] (%)	(15.0)				(17.5)				(16.8)			
ASSYM[7] (%)	(3.1)				(4.1)				(3.3)			
COLDSYM[8] (%)	(22.9)				(30.0)				(21.5)			

Notes:

1. Packs per day smoked by the mother
2. Family uses a gas stove
3. Years schooling of family head
4. Family size/number of rooms in house
5. Mother is head of household
6. Child has had bronchitis symptoms in last 3 years
7. Child has had asthma symptoms on day of lung function test
8. Child has had cold symptoms on day of lung function test

Table 2
Descriptive Statistics
Pollution Variables:

		Mean	Standard Deviation	Maximum	Pearson Correlation-Coefficients: RSP	SULF	NIT
Fall	TSP	74.1		156	.87	-.77	.95
	RSP	28.5		49		-.92	.97
N = 880	SULF	9.1		11			-.93
	SULF	1.5		3			
	TSP	73.6	32.1	157		-.34	.83
N = 1792	SULF	7.6	2.8	11			-.09
	NIT	1.1	0.8	4			
Winter	TSP	85.6	60.9	306	.75	.48	.35
	RSP	44.4		110		-.03	.56
N = 4303	SULF	7.4	3.25	14			-.25
	NIT	1.0	0.63	2			
Spring	TSP	64.4		252	.90	.05	.33
	RSP	22.1		121		.20	-.01
N = 2868	SULF	7.9		24			-.09
	NIT	1.5		2			
	TSP	68.0	31.2	252		-.07	.33
N = 4816	SULF	10.7	7.7	26			-.39
	NIT	1.3	0.5	2			

other), packs per day smoked by the mother, and, separately, by the father, an interaction term for age by sex by mother's smoking status, type of stove heat (gas or other), education of family head, number of family members per room, family size, mother head of household (yes or no), the presence of active asthma or cold symptoms on the test day, the presence of chronic bronchitis, and dummy variables for the teams and the spirometers used in the lung function tests. The latter variables were used to capture any biases in the administration of the tests or the measurement of test scores. Finally, we added pollution variables, including the 24-hour average concentrations of pollutants the day of the FEV exam and the day before the exam, these terms in squared or log form (to test for particular nonlinear relationships), the average daily values for the previous week, the maximum daily concentration for the previous week, and the average for the previous month.

We assume, for the first part of this analysis, that the presence (or probability) of asthma, chronic bronchitis, or cold symptoms on the one day per season a child was tested is unrelated to pollutant exposure or the other independent variables. These assumptions allow the full sample to be analyzed. Note, however, that symptom and chronic disease variables may be determined simultaneously with the lung function variable. In this case, attempts to explain any one of these dependent variables with OLS may lead to biased and inconsistent estimates of the regression coefficients (Pindyck and Rubinfield, 1976, Chapter IX). Because of limited time and budget, we leave this more complicated problem of simultaneous equation estimation for further research.

Because the assumption of no relationship between symptoms and. pollution is unrealistic, in the subsequent analysis, we estimate regressions for the subpopulation of normal, asymptomatic children and then the subpopulation of symptomatic children who participated in the Birmingham study. A symptomatic child is defined as one who had asthma symptoms or cold symptoms on the day of the test, or who has chronic bronchitis. An asymptomatic child is defined as one who has none of these.

Collinearity problems were examined with the aid of a diagnostic package in SAS based on Belsley, Kuh, and Welsch (1980). We found little collinearity between the non-pollution variables, except for age and height. The pollution variables were another story. As table 2 shows, TSP and RSP are highly correlated. Therefore, we ran separate regressions on each. As the results for TSP were superior, most of the regressions were run using this expression for particulates. Pollution variables for the other pollutants were included only when they were uncorrelated with the particulate measure. The daily pollution variables are sometimes correlated with their lagged values. (See Appendix Table A-1.) We tried alternative forms of the lagged variables both to find those which were uncorrelated and to explore those which fit the data best.

Finally, we made several attempts at fitting a Box-Tidwell specification to the data. This nonlinear specification is $FEV=a+b_i X_i^{\lambda} i+e$, where X_i is the ith independent variable. This specification allows the concentration-response function to be concave, convex, or a straight line $(\lambda=1)$, with this choice determined by the criterion of minimizing the sum of squared residuals. Unfortunately, with more than two OF three independent variables in the regression, a solution may not emerge. As such was the case here, we do not discuss this specification further.

Results--Full Sample

The results of the above analysis on symptomatic and asymptomatic children combined are presented in Tables 3 through 6. Table 3 presents representative results with pollution variables in linear form but not lagged.

The variables in this general model explain, in each of the three rounds, about 74% of the variation in FEV scores. Age, height, sex, and race account for most of this variation, along with dummy variables for presence of chronic bronchitis, asthma symptoms, and cold symptoms. The size of the age and sex coefficients are close to those of Hasselblad, et al. (1981). Of lesser importance, but with generally significant, positive effects on FEV is educational attainment. Family size and crowding in the house have unexpected positive, but small, effects on FEV scores. Both the team performing the tests and the spirometer used to measure the results were found to significantly affect FEV scores. Fortunately, because the testing protocol rotated teams and machines over all six exposure areas, the bias imparted to the FEV scores can be removed without affecting estimates of the effect of pollution exposures.

Finally, the TSP and sulfate, pollution variables are found to take the expected negative sign and to have significant effects on FEV. The TSP coefficient is quite stable and significant in fall and spring; the sulfate coefficient is highly significant in all rounds although its size varies by a factor of 5 over the three seasons. In fall and spring, nitrates had a negative but quite insignificant effect on pulmonary function scores (not shown), although high collinearity between nitrates and TSP in the fall may be obscuring such an effect. Surprisingly, in winter (shown as regression (3), nitrates appear to have a positive and significant effect on lung

Table 3
Regression Results: Linear Pollution Specifications

	Fall (1)	Winter (2)	(3)	Spring (4)	(5)
INTERCEP	-2.1138*	-2.4792*	-2.5110*	-2.5698*	-2.5590*
	(0.0849)	(0.0525)	(0.0509)	(0.0508)	(0.0514)
TSP	-0.0004*	-0.0001	-0.0003*	-0.0004*	-0.0003*
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SULF	-0.0118*	-0.0055*		-0.0025*	
	(0.0022)	(0.0010)		(0.0004)	
NITRATE			0.0312*		-0.0092
			(0.0057)		(0.0064)
AGE	0.0241*	0.0272*	0.0269*	0.0130*	0.0129*
	(0.0048)	(0.0030)	(0.0030)	(0.0029)	(0.0029)
HEIGHT	0.0599*	0.0634*	0.0634*	0.0683*	0.0683*
	(0.0018)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
BOYS	0.1059*	0.0893*	0.0892*	0.0857*	0.0857*
	(0.0113)	(0.0070)	(0.0070)	(0.0068)	(0.0068)
WHITE	0.1773*	0.1812*	0.1661*	0.2026*	0.2029*
	(0.0133)	(0.0099)	(0.0107)	(0.0085)	(0.0085)
TEAM1		-0.0055	0.0007	-0.0256*	-0.0258*
		(0.0090)	(0.0090)	(0.0090)	(0.0090)
TEAM2		-0.0483*	-0.0477*	-0.0569*	-0.0569*
		(0.0088)	(0.0088)	(0.0086)	(0.0086)
TEAM3		0.0280*	0.0220*	-0.0392*	-0.0396*
		(0.0096)	(0.0095)	(0.0082)	(0.0082)
SPRNO1	0.0217	0.0678*	0.0597*	0.0426*	0.0423*
	(0.0151)	(0.0094)	(0.0093)	(0.0087)	(0.0087)
SPRNO2	0.0617*	0.0563*	0.0532*	0.0441*	0.0446*
	(0.0139)	(0.0091)	(0.0090)	(0.0084)	(0.0084)
SPRNO3	-0.0011	-0.0005	0.0018	0.0640	0.0640*
	(0.0138)	(0.0091)	(0.0091)	(0.0084)	(0.0084)
PPDMOM	0.0113	-0.0072	-0.0058	-0.0128	-0.0127
	(0.0146)	(0.0087)	(0.0087)	(0.0085)	(0.0085)
PPDDAD	0.0025	0.0033	0.0030	0.0008	0.0007
	(0.0089)	(0.0051)	(0.0051)	(0.0051)	(0.0051)
GAS	-0.0033	0.0017	0.0018	0.0081	0.0079
	(0.0118)	(0.0079)	(0.0008)	(0.0076)	(0.0076)
EDU	-0.0007	0.0023*	0.0023*	0.0023*	0.0023*
	(0.0020)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
CROWD	0.0384*	0.0291*	0.0294*	0.0316*	0.0321*
	(0.0052)	(0.0100)	(0.0099)	(0.0097)	(.0097)
FAMSIZE	0.0030	0.0062*	0.0061*	0.0045	0.0046*
	(0.0032)	(0.0026)	(0.0026)	(0.0024)	(0.0024)
MOMHEAD	-0.0133	0.0096	0.0089	-0.0012	-0.0014
	(0.0133)	(0.0101)	(0.0101)	(0.0095)	(0.0095)
BRONC	-0.0124	0.0010	-0.0008	-0.0168*	-0.0168*
	(0.0138)	(0.0082)	(0.0082)	(0.0081)	(0.0081)

Table 3 (Continued)

	Fall	Winter		Spring	
	(1)	(2)	(3)	(4)	(5)
ASSYMP	-0.0972*	-0.0744*	-0.0720*	-0.0741*	-0.0736*
	(0.0281)	(0.0157)	(0.0157)	(0.0168)	(0.0168)
CDSYMP	-0.0283*	-0.0277*	-0.0270*	-0.0282*	-0.0279*
	(0.0116)	(0.0068)	(0.0068)	(0.0073)	(0.0073)
PKBYAG	-0.0036	-0.0020	-0.0020	-0.0004	-0.0004
	(0.0021)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
F	251.4	562.0	563.4	591.0	566.5
R^2 adj.	.7366	.7500	.7504	.7381	.7382
N	1792	4303	4303	4816	4816

Numbers in parenthesies are standard errors.
"*" means the coefficient is significant at the 95% level.

function. Although high collinearity can lead to such perversities, collinearity diagnostics reveal no particular problems with the nitrate variable. Given the instability of the results for nitrates across rounds and the insignificant effect the nitrate variable has on the results for other variables (except that its inclusion tends to increase the significance of TSP in winter), we drop this variable from further consideration. Further research on this anomolous effect appears warranted.

To obtain some idea of the size of these pollution effects, we computed elasticities at mean values (the percentage change in the mean FEV caused by a one percent change in the mean concentration of a pollutant) and compared the size of the pollution coefficients to those of other variables. The TSP elasticities range from -0.004 to -0.019, with most in the -0.014 to -0.019 range. Even at the high end of this range the effects are quite small. Sulfates appear to have somewhat larger elasticities--from -0.017 to -0.072. Considering regression (5), the effect of a 50% increase in either TSP or sulfates would be equivalent to being about one year younger or (since age and height are correlated) one-seventh of an inch shorter.

Table 4 presents the results of regressions with a number of different lagged specifications for the pollution variables. Most of those we tried evidenced too much collinearity between the day-of-test pollution variables and its lagged value to be useful. Lagged values that were significant usually took the reverse sign of the current day variable. When only lagged variables were included in the regressions, their coefficients were generally smaller and less significant than those for the current day.

Table 4
Regression Results: Alternative Lag Structures, By Round

	Fall (1)	Fall (2)	Winter (3)	Winter (4)	Spring (5)	Spring (6)	Spring (7)
TSP	-.00029 (.00021)	-.0011 (.00076)	.00027* (.00009)	-.00009 (.00007)	-.00038* (.00010)	-.00043*^{c} (.00017)^c	-.00026* (.00011)
TSPLG	.00020^c (.00032)		-.00035* (.00011)		.000009 (.0002)		
TSPWK		-.0011^c (.0010)				.00016^c (.00033)	
TSPMAX				.00010 (.00016)c			-.00031* (.00009)
SULF	-.0175*^{c} (.0068)	-.0057^c (.0063)	-.0095* (.0016)	-.0110^c (.0024)^c	-.0070*c (.0019)	-.0023* (.0006)	-.0029* (.0007)
SULFLG	.0033^c (.0038)		.0049* (.0011)		.0044* (.0019)^c		
SULFWK		-.0132*^{c} (.0130)				-.00023c (.0010)	
SULFMAX				.0048*^{c} (.0012)			.00005 (.00066)
F	228.4	228.5	519.9	519.4	544.7	543.5	545.3
R^2 adj.	.7364	.7365	.7510	.7508	.7384	.7380	.7386
N	1792	1792	4303	4303	4816	4816	4816

c = highly collinear with another pollution variable.
* = significant at the 95% level.
Numbers in parentheses are standard errors.

Table 5 reports the results for regressions identical to those in table 3 except that TSP and/or sulfate values are squared before the regressions are run. Using this squared specification has the effect of testing the plausible proposition that higher concentrations of pollution have a more than proportional negative effect on lung function.

The squared specification performs similarly to the linear form. Both pollutants show negative and significant effects in each season (except TSP in the winter under either specification). While on a priori grounds we may have reason to chose the squared specification over the linear, it is possible to test whether one is statistically superior to the other at explaining variation in lung function. We use the C-test of Davidson and MacKinnon. (1981), which involves running two regressions. In one regression, the dependent variable is the residuals from a regression run using the squared pollution variables (such as in table 5). The independent variable is a term equal to the difference in the predicted values for lung function over the squared and linear specifications. The second regression explains variation in the residuals for the specification with the linear pollution variables using the same independent variable. The coefficient on these differences is $\hat{\alpha}$. If the hypothesis $\alpha = 0$ cannot be rejected in the first regression but $\alpha = 0$ can be rejected in the second, the specification with the squared pollution variables is superior. The reverse outcome would indicate that the linear specification is superior.

The results of this test, shown below, do not indicate a clear winner. Thus, we conclude that both specifications explain variation in lung function equally well.

Table 5

Regression Results: Squared Pollution Specification; Marginal Effects of the Linear and Squared Specifications

	Fall	Winter	Spring
INTERCEP	-2.1743*	-2.4995*	-2.6060*
	(0.0815)	(0.0508)	(0.0500)
TSPSQ	-2.1×10^{-6}*	-9.2×10^{-8}	-1.4×10^{-6}*
	(7.9×10^{-7})	(2.1×10^{-7})	(4.3×10^{-7})
SULFSQ	-0.00075*	-0.00029*	-7.0×10^{-5}*
	(0.00015)	(5.9×10^{-5})	(1.3×10^{-5})
AGE	0.0240*	0.0271*	0.0130*
	(0.0047)	(0.0030)	(0.0030)
HEIGHT	0.0599*	0.0634*	0.0683*
	(0.0017)	(0.0011)	(0.0011)
BOYS	0.1061*	0.0893*	0.0858*
	(0.0113)	(0.0070)	(0.0068)
WHITE	0.1839*	0.1809*	0.2053*
	(0.0132)	(0.0101)	(0.0085)
TM1		-0.0064	-0.0275*
		(0.0091)	(0.0091)
TM2	--	-0.0484*	-0.0606*
		(0.0088)	(0.0085)
TM3		0.0280*	-0.0406*
		(0.0096)	(0.0082)
SPRNO1	0.0212	0.0681*	0.0426*
	(0.0152)	(0.0094)	(0.0087)
SPRNO2	0.0624*	0.0561*	0.0456*
	(0.0139)	(0.0091)	(0.0084)
SPRNO3	-0.0014	-0.00059	0.0640*
	(0.0138)	(0.0091)	(0.0084)
PPDMOM	0.0108	-0.0072	-0.0128
	(0.0146)	(0.0087)	(0.0085)
PPDDAD	0.0024	0.0034	0.00072
	(0.0089)	(0.0051)	(0.0051)
GAS	-0.0015	0.0014	0.0097
	(0.0117)	(0.0079)	(0.0076)
EDU	-0.00046	0.0022+	0.0026*
	(0.0020)	(0.0012)	(0.0012)
CROWD	0.0381*	0.0287*	0.0323*
	(0.0151)	(0.0099)	(0.0097)
FAMSIZE	0.0028	0.0061*	0.0045+
	(0.0032)	(0.0026)	(0.0024)
MOMHEAD	-0.0133	0.0096	-0.0014
	(0.0133)	(0.0101)	(0.0095)
BRONC	-0.0128	0.0013	-0.0174*
	(0.0138)	(0.0082)	(0.0081)
ASSYMP	-0.0981*	-0.0746*	-0.0747*
	(0.0281)	(0.0157)	(0.0169)
CDSYMP	-0.0285*	-0.0276*	-0.0280*
	(0.0116)	(0.0068)	(0.0073)
PKBYAG	-0.0035+	-0.0020+	-0.00036
	(0.0021)	(0.0011)	(0.0011)
Adj R^2	0.7364	0.7501	0.7375
F	251.2	562.3	589.2
N	1792	4303	4816

$\partial FEV/\partial X$; X =	Fall TSP	Fall Sulfates	Winter TSP	Winter Sulfates	Spring TSP	Spring Sulfates
$\hat{\beta}$ (linear)	-.0004	-.0118	-1.5×10^{-5}	-.0053	-.0004	-.0025
$2\hat{\beta}\bar{X}$ (squared)	-.0003	-.0115	-1.6×10^{-5}	-.0043	-.0002	-.0015
$x^* = \beta_L/2\beta_s$	100	7.82	83.3	9.0	143	17.86
Max X	157	11	306	14	252	26
Mean X	73.6	7.6	85.6	7.4	68	10.7

Numbers in parentheses are standard errors.
"*" means the coefficient is significant at the 95% level.
"+" means the coefficient is significant at the 90% level.

Model	Estimated Value of α	t Value	H_0 ($\alpha = 0$)
$a + bX$	1.55	1.070	accept
$a + bX^2$	-0.55	-0.381	accept

The magnitude of the effect of each pollutant on lung function can be compared for the two alternative specifications and for each season. At the bottom of table 5 the marginal effects are presented (evaluated at the mean pollutant concentration in the case of the squared specification). In all three seasons and for both pollutants, the marginal effects are larger for the linear specification. However, at higher pollution levels, the marginal effects increase using the squared specification. Indeed, as shown in table 5, at concentrations that exceed the mean but are, nevertheless, commonly observed in each season, the marginal effects on lung function estimated with the squared specification exceed those estimated with the linear.

The results can also be compared for a given specification, but across seasons. The marginal TSP effects are quite similar across seasons with either the linear or the squared specification. However, the marginal sulfate effects are greatest in the fall and lowest in the spring.[3]

---

3. Assuming the squared relationship is the correct one, this finding, is, at first glance, surprising. The squared specification predicts that lung function.readings will be more than proportionally lower for those exposed to higher sulfate concentrations. Because the marginal effect of sulfates on lung function is estimated to be greatest in the fall and smallest in the spring, one would then expect to find that average pollution values in the fall exceed those in the spring (assuming that the distributions of sulfate concentrations are similar in both seasons). However, contrary to expectations, table 5 reveals that average sulfate concentrations are much lower in the fall than in the spring. Nevertheless, when the air quality data are transformed into their more appropriate log-normal distribution, the apparent contradiction disappears. Comparing averages now, we find that the mean log sulfate concentration in the fall is significantly <u>greater</u> than that in the spring.

Finally, we investigated two additional specifications for the pollution variables: 1) the log form and 2) the linear form with a term for the interaction between TSP and sulfates. Regressing the log of pollution concentrations against lung function scores has the effect of testing for a less than proportional relationship (diminishing marginal effect) of air pollution on lung function. Table 6, which summarizes the results of the linear, squared, log, and interaction specifications, shows that the log form is no better or worse at explaining variation in lung function than the linear and squared specifications. As the underlying concentration-response function implied by the log form is less medically plausible than the linear and squared form, we say no more about it. Table 6 also shows that adding an interaction term to the linear specification does not significantly improve fit. Collinearity diagnostics (Belsley, Kuh, and Welsch, 1980) reveal high collinearity among the three pollution variables (TSP, SULFATES, and TSP* SULFATES), ruling out a definitive statement about the existence of interaction effects. As expected when collinearity is present, the linear pollution variables become less significant.

Results on Asymptomatics-Symptomatics

The next part of the analysis concerns the differential effects of pollution (and other variables) on symptomatic vs. asymptomatic children. The results are presented in table 7.

By comparing the size of the comparable and significant pollution coefficients for the asymptomatic and symptomatic regressions [(1) vs. (4), (2) vs. (5), (3) vs. (61], it is apparent that pollution has a greater

Table 6

Regression Results: Alternative Pollution Specifications, by Round

	FALL			WINTER				SPRING				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
TSP	-.0004* (.0002)			-.0001 (.0001)			-.00030 (.00031)	-.0004* (.0001)			-.00067+ (.00037)	
TSPSQ		-2.1×10^{-6}* (7.87×10^{-7})			-3.5×10^{-8} (1.2×10^{-6})				-1.4×10^{-6}* (4.3×10^{-7})			
LOGTSP			-.0250+ (.0148)			-.0139+ (.0074)				-.0249* (.0081)		
SULF	-.0118* (.0022)			-.0055* (.0010)			-.0021 (.0024)	-.0025* (.0004)			-.0039* (.0017)	
SULFSQ		-.00075* (.00015)			-.00030* (5.7×10^{-5})				-7.0×10^{-5}* (1.3×10^{-5})			
LOGSULF			-.0810* (.0146)			-.0406* (.0081)				-.0288* (.0043)		
TSP*SULF							-.00007+ (.00004)				2.7×10^{-5} (3.1×10^{-5})	-3.8×10^{-5}* (5.8×10^{-6})
F	251.4	251.2	251.5	562.0	562.3	561.7	538.9	591.0	589.2	591.6	566.4	617.0
R^2	0.7366	0.7364	0.7366	0.7500	0.7500	0.7499	0.7500	0.7381	0.7375	0.7383	0.7381	0.7378
N	1792	1792	1792	4303	4303	4303	4303	4816	4816	4816	4816	4816

Note: TSPSQ = the square of total suspended particulate concentrations
LOGTSP = the log of total suspended particulate concentrations
SULFSQ = the square of sulfate concentrations
LOGSULF = the log of sulfate concentrations
TSP*SULF = total suspended particulate concentration times sulfate concentration
"*" means coefficient is significant at the 95% level.
"+" means coefficient is significant at the 90% level.
Numbers in parentheses are standard errors.

Table 7
Regression Results: Asymptomatics vs. Symptomatics

	Asymptomatics			Symptomatics		
	(1) Fall	(2) Winter	(3) Spring	(4) Fall	(5) Winter	(6) Spring
Intercept	-2.0570*	-2.4305*	-2.5955*	-2.2752*	-2.5433*	-2.5519*
	(0.1032)	(0.0702)	(0.0622)	(0.1487)	(0.0799)	(0.0886)
TSP	-0.0004*	0.0002	-0.0002*	-0.0004	-0.0006*	-0.0006*
	(0.0002)	(0.0002)	(0.0001)	(0.0003)	(0.0002)	(0.0002)
SULF	-0.0072*	-0.0069*	-0.0022*	-0.0207*	-0.0036*	-0.0029*
	(0.0027)	(0.0014)	(0.0005)	(0.0036)	(0.0016)	(0.0006)
AGE1	0.0260*	0.0303*	0.0137*	0.0201*	0.0227*	0.0103*
	(0.0056)	(0.0041)	(0.0036)	(0.0089)	(0.0046)	(0.0051)
HEIGHT	0.0582*	0.0623*	0.0684*	0.0637*	0.0647*	0.0684*
	(0.0021)	(0.0015)	(0.0014)	(0.0033)	(0.0017)	(0.0019)
BOYS	0.1024*	0.0969*	0.0922*	0.1216*	0.0836*	0.0749*
	(0.0139)	(0.0094)	(0.0084)	(0.0194)	(0.0106)	(0.0118)
WHITE	0.1939*	0.1853*	0.2037*	0.1642*	0.1764*	0.2031*
	(0.0167)	(0.0129)	(0.0103)	(0.0222)	(0.0155)	(0.0153)
TEAM1		-0.0013	-0.0196		-0.0103	-0.0332*
		(-0.0122)	(0.0111)		(0.0135)	(0.0156)
TEAM2		-0.0508*	-0.0552*		-0.0459*	-0.0573*
		(0.0120)	(0.0105)		(0.0132)	(0.0150)
TEAM3		0.0278*	-0.0363*		0.0206	-0.0426*
		(0.0131)	(0.0100)		(0.0139)	(0.0142)
SPRNO1	0.0322	0.0635*	0.0399*	0.0005	(0.0700)	0.0448*
	(0.0183)	(0.0125)	(0.0108)	(0.0270)	(0.0146)	(0.0150)
SPRNO2	0.0646*	0.0532*	0.0377*	0.0552*	0.0553*	0.0546*
	(0.0172)	(0.0120)	(0.0103)	(0.0238)	(0.0141)	(0.0144)
SPRNO3	-0.0198	0.0047	0.0604*	0.0276	-0.0092	0.0688*
	(0.0176)	(0.0122)	(0.0104)	(0.0224)	(0.0139)	(0.0146)
PPDMOM	0.0093	-0.0000	-0.0063	0.0213	-0.0116	-0.0221
	(0.0181)	(0.0120)	(0.0106)	(0.0249)	(0.0128)	(0.0145)
PPDDAD	-0.0011	-0.0008	-0.0024	-0.0014	0.0055	0.0053
	(0.0112)	(0.0070)	(0.0063)	(0.0146)	(0.0076)	(0.0085)
GAS	-0.0131	-0.0055	0.0120	0.0262	0.0113	0.0013
	(0.0143)	(0.0105)	(0.0093)	(0.0206)	(0.0121)	(0.0132)
EDUC	-0.0010	0.0016	0.0016	-0.0005	0.0037*	0.0033
	(0.0024)	(0.0016)	(0.0014)	(0.0034)	(0.0019)	(0.0020)
CROWD	0.0290	0.0195	0.0373*	0.0491*	0.0371*	0.0223
	(0.0189)	(0.0136)	(0.0120)	(0.0258)	(0.0148)	(0.0168)
FAMSIZE	0.0030	0.0048	0.0048	0.0040	0.0068	0.0031
	(0.0038)	(0.0034)	(0.0028)	(0.0059)	(0.0041)	(0.0045)
MOMHEAD	-0.0268	0.0097	-0.0014	0.0111	0.0032	-0.0033
	(0.0161)	(0.0136)	(0.0115)	(0.0239)	(0.0153)	(0.0169)
PKBYAG	-0.0040	-0.0031*	-0.0014	-0.0041	-0.0012	0.0013
	(0.0025)	(0.0015)	(0.0014)	(0.0037)	(0.0017)	(0.0020)

Table 7 (Continued)

	Asymptomatics			Symptomatics		
	(1) Fall	(2) Winter	(3) Spring	(1) Fall	(2) Winter	(3) Spring
F	194.3	357.1	448.4	101.3	283.7	225.4
R^2 adj.	0.7426	0.7488	0.7450	0.7236	0.7473	0.7192
N	1140	2390	3063	652	1913	1753

Numbers in parentheses are standard errors.
"*" means the coefficient is significant at the 95% level.

effect on the FEV of symptomatic children than on the FEV of asymptomatic children. The only exception is the sulfate variable in (2) vs. (5). However, the differences between comparable coefficients are significant at the 95% level only for sulfates in the fall round. In this instance, the FEV-pollution elasticity (the percentage change in mean FEV scores from a 1% change in the mean sulfate concentration) for symptomatics is -0.108, which means that the symptomatic reaction is considerably greater than that of the general population (-0.072).

Results for some of the other independent variables are interesting in comparison to those for the combined sample. Crowding, education, and family size are all less robust variables in the split samples. In addition, crowding, at least in the fall and winter rounds, appears to be more beneficial to symptomatics than asymptomatics.

The Consequences of Increasing Pollution Data Detail

Our data base contains pollution data corresponding to the month and the day, respectively, that the lung function tests were given. Thus, we are in a position to examine the effect of using more or less finely detailed pollution data on explaining differences in lung function. While clinical research clearly identifies a more or less immediate reaction of FEV to hourly exposures (followed perhaps by adaptation), the response of FEV to longer-term exposures has not been as well established. By replacing daily concentrations with corresponding monthly values in the above regressions, we are in a position to compare pollution coefficients at two levels of temporal detail. Assuming that longer-term effects are zero, then the monthly data are just imperfect proxies of daily data. In this case we would expect the monthly variables to be less significant and have a smaller coefficient than the dally variables; That this is so is intuitively reasonable. In addition, if using the monthly concentration is thought of as measuring the daily concentration with error, the estimate of the FEV pollution relationship will be biased and inconsistent. In general, the true relationship will be underestimated (Johnston, 1972, p. 281-3).

An alternative hypothesis is that longer-term (monthly) exposures affect lung function more than daily exposures. In this case the monthly correlation variable should be more significant and have a smaller (larger) negative coefficient than the daily variable. Mixed results would, of course, leave this issue in doubt.

To test the alternative hypotheses we re-ran most of the above regressions, including those with both the full and the split samples, with monthly pollution values as explanatory variables. Table 8 shows the results.

Although only 18 comparisons of pollution coefficients were made (two pollutants, three seasons; full sample, symptomatics, and asymptomatics), the results are striking. There are no instances where the coefficient on the monthly pollution variable is significant while the daily coefficient is not. This is evidence that the effect of pollution on FEV is a very short-term phenomenon. Note also that in all 18 cases r^2 and F statistics for the regressions with daily variables exceed those using monthly variables; on average $r^2_D/r^2_M = 1.0037$ and $F_D/F_M = 1.0133$. This result holds even for the cases where both monthly and daily pollution variables are significant.

Table 8

Comparison of Regression Results

Monthly (M) versus Daily (D) Pollution Coefficients

	M,* D*[1]		D*, M		M*, D		M, D		Totals	
	TSP	Sulfates	TSP	Sulfates	TSP	Sulfates	TSP	Sulfates	TSP	Sulfates
|M|>|D|	1[2]	5	0	0	0	0	1	0	2	5
|M|<|D|	1	0	4	4	0	0	2	0	7	4
Totals	2	5	4	4	0	0	3	0	9	9

1. An "*" means that the pollution coefficient is significant at the 95% level.

2. Each entry in the table refers to comparisons in the significance and size of the monthly vs. the daily pollution coefficients estimated in otherwise identical regression specifications. In column 1, of two such comparisons where both the monthly and daily TSP coefficients were significant, in one the absolute value of the monthly coefficient exceeded that of the daily value, in the other the reverse was true.

Viewing the two pollution variables separately, the picture becomes a bit blurred. The above result quite clearly applies to TSP, where in 5 of 6 cases the daily coefficients are larger (in absolute value) than the monthly coefficients. However, for sulfates, only 4 of 9 cases show this relationship. Further, the coefficient for the monthly variable is always larger than that for the daily variable whenever each measure is significantly related to FEV scores. We have no explanation for this result.

Conclusion

We found statistically significant linear relationships between FEV.75 measurements for a sample of Birmingham, Alabama school children taken during three consecutive seasons of the 1972-73 school year and the 24-hour average concentrations of TSP (or RSP) and sulfates on the day of the lung function tests. Because of high collinearity between TSP and RSP the separate effects of each cannot be determined. Nitrates had a generally insignificant effect, except in the winter round when exposure to higher levels of nitrates appears to have increased lung function scores. Attempts to isolate the effect of lagged pollution exposures on FEV scores were generally unsuccessful. In addition, alternative nonlinear specifications of the lung function-pollution relationship performed no better than the linear specification. We also found that daily pollution data are generally better at explaining lung function scores than monthly average data. Therefore, using monthly instead of daily pollution data are likely to cause an error-in-variables problem. Considering the other independent variables in the study, we found correctable biases introduced by the teams and machines used to conduct the tests and a positive relationship between the head of household's educational attainment and the child's FEV score. However, neither smoking habits nor type of stove fuel used had any effect. Finally, we found unexpectedly that more crowding in the home and larger family sizes were associated with higher lung function scores.

References

Belsley, A., E. Kuh, R. E. Welsch. 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity (New York, John Wiley & Sons).

Davidson, R. and J.G. MacKinnon. 1982. "Several Tests for Model Specifications in the Presence of Alternative Hypothesis," Econometrica vol. 49, no. 3 (May).

Dockery, D. W., N. R. Cook, B. G. Ferris, Jr., F. E. Speizer, J. D. Spengler, and J. H. Ware, 1981. "Change in Pulmonary Function in Children Associated with Air Pollution Episodes," presented at the 74th Annual Meeting of the Air Pollution Control Association, Philadelphia, Pa. (June).

Hasselblad, V., C. G. Humble, M. G. Graham, and H. S. Anderson. 1981. "Indoor Environmental Determinants of Lung Function in Children," American Review of Respiratory Disease vol. 123, No. 5, May.

Johnston, J. 1972. Econometric Methods (New York, McGraw-Hill, Inc.)

Krupnick, A. and W. Harrington. 1984. Air Pollution and Lung Function in Children, prepared for Economic Analysis Branch, OPPE, USEPA (September).

Lawther, P.J., R.E. Waller, and M. Henderson. 1970. "Air Pollution and Exacerbations of Bronchitis," Thorax 25: 525-539.

Lawther, P.J., A.G.F, Brooks, P.W. Lord, and R.E. Waller. 1974. "Day-to-Day Changes in Ventilatory Function in Relation to the Environment. Part I. Spirometric Values," Environmental Research 7: 24-40.

Lunn, J.E., J. Knowelden, and A.J. Handyside. 1967. "Patterns of Respiratory Illness in Sheffield Infant School children," British Journal of Prev. Social Medicine 2l: 7-16.

_______, ________, and J.W. Roe. 7970. "Patterns of Respiratory Illness in Sheffield Junior School Children: A Follow-up Study," British Journal of Prev: Social Medicine 24: 223.228.

Pindyck, R. S. and D. L. Rubinfeld. 1976. Econometric Models and Economic Forecasts (New York, McGraw-Hill, Inc.).

Table A-1
Correlation Coefficients Between Pollution Variables, by Round

Fall (n = 1792)

	TSP	SULF	NITR	TSPLG1	SULFLG1	TSPWK	SULFWK	TSPMAX	SULFMAX	TSPMO	SULFMO
TSP[1]	--	-.34	.83	.22	-.53	-.50	-.84	-.49	-.83	-.64	-.81
SULF			-.09	.42	.83	.49	.71	.60	.75	.48	.68
NITR				.67	-.44	.04	-.76	.05	-.51	-.24	-.47
TSPLG[1]					-.02	.70	-.14	.72	.24	.47	.28
SULFG[1]						.20	.79	.31	.67	.18	.53
TSPWK							.37	.99	.74	.97	.90
SULFWK								.42	.90	.58	.83
TSPMAX									.78	.88	.81
SULFMAX										.84	.98
TSPMO											.92
SULFMO											--

Winter (n = 4303)

	TSP	SULF	NITR	TSPLG1	SULFLG1	TSPWK	SULFWK	TSPMAX	SULFMAX	TSPMO	SULFMO
TSP	--	.48	.35	.55	-.15	.57	-.37	.66	-.07	.76	.04
SULF			-.25	.49	.52	.44	neg	.78	.59	.61	.58
NITR				.30	.02	-.06	-.06	-.21	-.12	.13	-.18
TSPLG[1]					.42	.41	-.22	.64	.06	.58	.05
SULFLG[1]						.08	.49	.15	.76	.09	.61
TSPWK							-.01	.67	.29	.67	.35
SULFWK								-.42	.71	-.54	.73
TSPMAX									.17	.84	.20
SULFMAX										.05	.91
TSPMO											.03
SULFMO											--

Table A-1 (Continued)

<u>Spring</u> (n = 4816)

	TSP	SULF	NITR	TSPLG1	SULFLG1	TSPWK	SULFWK	TSPMAX	SULFMAX	TSPMO	SULFMO
TSP	--	-.07	.33	.60	-.10	.76	.16	.38	.04	.66	.08
SULF			-.39	-.09	.98	-.25	.47	-.26	.79	-.26	.93
NITR				.21	-.51	.10	-.20	.15	-.42	.03	-.44
TSPLG[1]					-.05	.72	.27	.42	.03	.60	.03
SULFLG[1]						-.20	.48	-.26	.78	-.25	.93
TSPWK							.38	.52	.07	.82	-.07
SULFWK								.13	.67	.17	.52
TSPMAX									-.04	.72	-.14
SULFMAX										.07	.75
TSPMO											-.03
SULFMO											--

<u>Notes:</u>

1. TSP, SULF, NITR (24-hour reading the day of the test); TSPLG1, SULFLG1 (24-hour reading the day before the test); TSPWK, SULFWK (average of daily 24-hour readings taken the week before the test day); TSPMAX, SULFMAX (the maximum daily readings during the week before the test date); TSPMO, SULFMO (the monthly average of the 24-hour readings taken before the test date).